



On matrix equations $X - AXF = C$ and $X - A\bar{X}F = C$

Ai-Guo Wu^{a,*}, Hao-Qian Wang^b, Guang-Ren Duan^c

^a Harbin Institute of Technology Shenzhen Graduate School, University Town of Shenzhen, Shenzhen 518055, PR China

^b Graduate School at Shenzhen, Tsinghua University, University Town of Shenzhen, Shenzhen 518055, PR China

^c Center for Control Theory and Guidance Technology, Harbin Institute of Technology, Harbin 150001, PR China

ARTICLE INFO

Article history:

Received 17 March 2007

Received in revised form 5 September 2008

Keywords:

Kronecker map

Real representation

Controllability matrix

Observability matrix

ABSTRACT

With the help of the concept of Kronecker map, an explicit solution for the matrix equation $X - AXF = C$ is established. This solution is neatly expressed by a symmetric operator matrix, a controllability matrix and an observability matrix. In addition, the matrix equation $X - A\bar{X}F = C$ is also studied. An explicit solution for this matrix equation is also proposed by means of the real representation of a complex matrix. This solution is neatly expressed by a symmetric operator matrix, two controllability matrices and two observability matrices.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The matrix equations $XF - AX = C$ and $X - AXF = C$ play important roles in theories and applications of stability and control. When $F = A^H$ (the conjugate transpose of A), the matrix equations are the well-known Lyapunov equation and Stein equation (or, discrete-time Lyapunov equation), respectively. There exist several numerical algorithms to compute the solutions of the above matrix equations such as the Bartels–Stewart algorithm [1–3], and the well-known Hessenberg–Schur algorithm [4]. However, for several applications it is important to obtain symbolic solutions of such matrix equations. In [5] the solutions for the Lyapunov equation and Stein equation, which are particularly suitable for symbolic implementation, are proposed for the case where the matrix F is in a companion form. For the matrix equation $XF - AX = C$ with the matrices A and F being both Jordan forms, Ma provided an explicit solution in the forms of finite double matrix series [6]. Besides, when the matrix F is in an arbitrary form, some explicit solutions for the aforementioned matrix equations are also given. Perhaps the well-known ones are the integral forms in [7] of the unique solution to the Lyapunov matrix equation, and the matrix power form in [8] of the unique solution to the Stein matrix equation. For the matrix equation $XF - AX = C$, some explicit solutions were established in terms of controllability matrix and observability matrix in [9,10]. Besides, in [11–14] some solutions for $XF - AX = C$ and $X - AXF = C$ are given according to the coefficient of the characteristic polynomial of matrix A or the so-called Faddev algorithm. In addition, the matrix equation $X - A\bar{X}F = C$ was investigated in [14], where \bar{X} denotes the conjugate of the complex matrix X . The existence of a solution to the matrix equation $X - A\bar{X}F = C$ was characterized, and the solution was derived in explicit form by means of real representations.

In this paper, we study the solutions of the matrix equations $X - AXF = C$ and $X - A\bar{X}F = C$. With the help of the so-called Kronecker map, some explicit solutions to the matrix equation $X - AXF = C$ are established, one of which is exactly the one given in [14]. Among these solutions, one is neatly represented in terms of a symmetric operator matrix, a controllability matrix and an observability matrix. In addition, we also derive some explicit solutions of the matrix equation $X - A\bar{X}F = C$ by means of the real representation proposed in [14] with the help of the proposed solutions to the matrix

* Corresponding address: Room 417, Building D, HIT Campus, University Town of Shenzhen, Shenzhen 518055, PR China. Tel.: +86 755 26033899.

E-mail address: ag.wu@163.com (A.-G. Wu).

equation $X - AXF = C$. Compared with the one given in [14], these solutions can be obtained from the original matrices, instead of the real representation matrices of the original matrices. One of these solutions is neatly expressed by a symmetric operator matrix, two controllability matrix and two observability matrices.

2. Preliminaries

2.1. Notations

In this paper, we use $\det(A)$, A^T , \bar{X} , $\text{trace}(A)$ and $\text{adj}(A)$ to denote the determinant, the transpose, the conjugate, the trace and the adjoint of the matrix A , respectively, and use $\lambda(A)$ to denote the set of eigenvalues of A . For the square matrix pair (E, A) , $\lambda(E, A)$ denotes the set of finite eigenvalues, that is, $\lambda(E, A) = \{s; \det(sE - A) = 0, s \text{ finite}\}$. Further, let $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times r}$, and $C \in \mathbb{C}^{m \times n}$, we have the following notations associated with these matrices:

$$Q_c(A, B) = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}$$

$$Q_o(A, C, k) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix},$$

$$f_{(I,A)}(s) = \det(I - sA) = \alpha_n s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + 1,$$

$$S_r(I, A) = \begin{bmatrix} I_r & \alpha_1 I_r & \alpha_2 I_r & \cdots & \alpha_{n-1} I_r \\ & I_r & \alpha_1 I_r & \cdots & \alpha_{n-2} I_r \\ & & \ddots & \ddots & \vdots \\ & & & I_r & \alpha_1 I_r \\ & & & & I_r \end{bmatrix}.$$

Obviously, $Q_c(A, B)$ is the controllability matrix of the matrix pair (A, B) , $Q_o(A, C, n)$ is the observability matrix of the matrix pair (A, C) and $S_r(I, A)$ is a symmetric operator matrix.

2.2. Kronecker map

In this subsection, we introduce the concept of the so-called Kronecker map. It is to be used in the next section.

Definition 1 ([15]). Let $T(s) = \sum_{i=0}^t T_i s^i \in \mathbb{C}^{m \times q}[s]$, $F \in \mathbb{C}^{p \times p}$ and $Z \in \mathbb{C}^{q \times p}$. The following matrix sum

$$\text{Syl}(T(s), F, Z) = \sum_{i=0}^t T_i Z F^i$$

is called Sylvester sum associated with $T(s)$, F and Z . For a fixed matrix $F \in \mathbb{C}^{p \times p}$, the Kronecker map is defined as

$$\mathcal{F}[T(s)] = \sum_{i=0}^t (F^T)^i \otimes T_i.$$

Based on the newly defined Kronecker map, we have

$$\text{vec}(\text{Syl}(T(s), F, Z)) = \mathcal{F}[T(s)] \text{vec}(Z).$$

The so-called Kronecker map possesses the following good property.

Lemma 1 ([15]). Let $X(s) \in \mathbb{C}^{q \times r}[s]$, $Y(s) \in \mathbb{C}^{r \times m}[s]$, then

$$\mathcal{F}[X(s)Y(s)] = \mathcal{F}[X(s)]\mathcal{F}[Y(s)].$$

Based on this property, we have the following conclusion.

Lemma 2. Let $T(s) \in \mathbb{C}^{q \times q}[s]$. Then

$$\mathcal{F}[\text{adj}T(s)]\mathcal{F}[T(s)] = \mathcal{F}[I_q \det T(s)].$$

Proof. From Lemma 1, we have

$$\begin{aligned} (\text{adj}T(s))T(s) &= I_q \det T(s) \\ \implies \mathcal{F}[(\text{adj}T(s))T(s)] &= \mathcal{F}[I_q \det T(s)] \\ \implies \mathcal{F}[\text{adj}T(s)]\mathcal{F}[T(s)] &= \mathcal{F}[I_q \det T(s)]. \end{aligned}$$

The proof is completed. ■

2.3. Real representation

Let $A \in \mathbb{C}^{m \times n}$, then A can be uniquely written as $A = A_1 + A_2 i$, $A_1, A_2 \in \mathbb{R}^{m \times n}$, $i = \sqrt{-1}$. Define a real representation σ as

$$A_\sigma = \begin{bmatrix} A_1 & A_2 \\ A_2 & -A_1 \end{bmatrix} \in \mathbb{R}^{2m \times 2n},$$

A_σ is called the real representation of the matrix A . This concept was firstly proposed in [14].

For an $n \times n$ complex matrix A , define $A_\sigma^i = (A_\sigma)^i$, and

$$P_j = \begin{bmatrix} I_j & 0 \\ 0 & -I_j \end{bmatrix}, \quad Q_j = \begin{bmatrix} 0 & I_j \\ -I_j & 0 \end{bmatrix},$$

where I_j is the $j \times j$ identity matrix. The real representation possesses the following property, which can be found in [14].

Lemma 3 (The properties of the real representation).

(1) If $A, B \in \mathbb{C}^{m \times n}$, $a \in \mathbb{R}$, then

$$\begin{cases} (A+B)_\sigma = A_\sigma + B_\sigma \\ (aA)_\sigma = aA_\sigma \\ P_m A_\sigma P_n = \bar{A}_\sigma. \end{cases}$$

(2) If $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{n \times r}$, $C \in \mathbb{C}^{r \times p}$, then

$$\begin{cases} (AB)_\sigma = A_\sigma P_n B_\sigma = A_\sigma (\bar{B})_\sigma P_r \\ (ABC)_\sigma = A_\sigma \bar{B}_\sigma C_\sigma \end{cases}.$$

(3) If $A \in \mathbb{C}^{n \times n}$, then A is nonsingular if and only if A_σ is nonsingular.

(4) If $A \in \mathbb{C}^{n \times n}$, then $A_\sigma^{2k} = ((AA^k)_\sigma P_n)$.

(5) If $A \in \mathbb{C}^{m \times n}$, then $Q_n A_\sigma Q_n = A_\sigma$.

Regarding the characteristic polynomial of the real representation of a complex matrix, we have the following theorem. This result has been given in [16]. For convenience, the proof is given in the Appendix. This theorem will play a vital role in solving the matrix equation $X - AXF = C$.

Theorem 1 ([16]). Let $A \in \mathbb{C}^{n \times n}$, then $f_{(I, A_\sigma)}(s) = f_{(I, A\bar{A})}(s^2) = f_{(I, \bar{A}A)}(s^2) \in \mathbb{R}[s]$.

Proof. The proof is provided in the Appendix. ■

With the above theorem, the following conclusion is immediately obtained in view of the result in [14].

Lemma 4. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$. Then

$$f_{(I, A_\sigma)}(F_\sigma) = (f_{(I, A\bar{A})}(F\bar{F}))_\sigma P_p.$$

3. Matrix equation $X - AXF = C$

In this section, we consider the following matrix equation

$$X - AXF = C, \tag{3.1}$$

where $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. Regarding the solution to this matrix equation, we have the following result.

Theorem 2. Given $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$, let

$$f_{(I, A)}(s) = \det(I - sA) = \alpha_n s^n + \cdots + \alpha_1 s + \alpha_0, \quad \alpha_0 = 1, \tag{3.2}$$

and

$$\text{adj}(I - sA) = R_{n-1}s^{n-1} + \cdots + R_1 s + R_0. \tag{3.3}$$

Then

(1) If X is a solution of matrix equation (3.1), then

$$X f_{(I, A)}(F) = \sum_{i=0}^{n-1} R_i C F^i.$$

(2) If X is the unique solution of matrix equation (3.1), then

$$X = \left(\sum_{i=0}^{n-1} R_i C F^i \right) f_{(I,A)}^{-1}(F).$$

Proof. With the concept of the Sylvester sum, matrix equation (3.1) can be written as

$$\text{Syl}(I - sA, F, X) = C.$$

It is equivalent to

$$\mathcal{F}[I - sA] \text{vec}(X) = \text{vec}(C).$$

By Lemma 2, we have

$$\mathcal{F}[I_p \det(I - sA)] \text{vec}(X) = \mathcal{F}[\text{adj}(I - sA)] \text{vec}(C).$$

So we have

$$\text{Syl}(I_p \det(I - sA), F, X) = \text{Syl}(\text{adj}(I - sA), F, C).$$

By simple computations, we have

$$\text{Syl}(I_p \det(I - sA), F, X) = X f_A(F),$$

and

$$\text{Syl}(\text{adj}(I - sA), F, C) = \sum_{i=0}^{n-1} R_i C F^i.$$

So the conclusions hold. ■

This theorem provides a very neat explicit solution to matrix equation (3.1). To obtain this solution, one needs the coefficients α_i , $i = 0, 1, \dots, n-1$, of the characteristic polynomial of the matrix pair (I, A) , and the coefficient matrices R_i , $i = 0, 1, \dots, n-1$, of the adjoint matrix $\text{adj}(I - sA)$. They can be obtained using the following so-called generalized Leverrier algorithm [17]:

$$\begin{cases} R_i = R_{i-1}A + \alpha_i I, & R_0 = I \\ \alpha_i = \frac{\text{tr}(R_{i-1}A)}{i}, & \alpha_0 = 1, \end{cases} \quad i = 1, 2, \dots, n. \quad (3.4)$$

In the following, we provide an equivalent statement of the above theorem.

Theorem 3. Given $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$, let

$$f_{(I,A)}(s) = \det(I - sA) = \alpha_n s^n + \dots + \alpha_1 s + \alpha_0, \quad \alpha_0 = 1.$$

Then

(1) If X is a solution of matrix equation (3.1), then

$$X f_{(I,A)}(F) = \sum_{k=0}^{n-1} \sum_{j=k}^{n-1} \alpha_k A^{j-k} C F^j. \quad (3.5)$$

(2) If X is the unique solution of matrix equation (3.1), then

$$X = \left(\sum_{k=0}^{n-1} \sum_{j=k}^{n-1} \alpha_k A^{j-k} C F^j \right) f_{(I,A)}^{-1}(F).$$

Proof. From (3.4), it is easily obtained

$$\begin{cases} R_0 = I \\ R_1 = \alpha_1 I + A \\ R_2 = \alpha_2 I + \alpha_1 A + A^2 \\ \dots \\ R_{n-1} = \alpha_{n-1} I + \alpha_{n-2} A + \dots + A^{n-1}. \end{cases} \quad (3.6)$$

This relation can be compactly expressed by

$$R_j = \sum_{k=0}^j \alpha_k A^{j-k}, \quad \alpha_0 = 1.$$

Thus we have

$$\begin{aligned} \sum_{j=0}^{n-1} R_j C F^j &= \sum_{j=0}^{n-1} \left(\sum_{k=0}^j \alpha_k A^{j-k} \right) C F^j \\ &= \sum_{k=0}^{n-1} \sum_{j=k}^{n-1} \alpha_k A^{j-k} C F^j. \end{aligned}$$

Combining this with Theorem 2, we complete the proof. ■

Remark 1. For a matrix $A \in \mathbb{C}^{n \times n}$, if $\det(sI - A) = \sum_{i=0}^n \alpha_i s^i$, then $\det(I - sA) = \sum_{i=0}^n \alpha_i s^{n-i}$. Based on this fact, it is easily found that expression (3.5) is exactly the first expression in Theorem 2.5 of [14].

Without loss of generality, now we assume that the matrix C in matrix equation (3.1) has the following decomposition

$$C = C_1 C_2, \quad C_1 \in \mathbb{C}^{n \times r}, \quad C_2 \in \mathbb{C}^{r \times p}, \quad (3.7)$$

where $r = \text{rank } C$. This can be generally obtained by applying the full-rank factorization. In this case, for the solution of matrix equation (3.1) we have the following results.

Theorem 4. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$, and the matrix C admits the decomposition (3.7). Then

(1) If X is a solution of (3.1), then

$$X f_{(I,A)}(F) = Q_c(A, C_1) S_r(I, A) Q_o(F, C_2, n).$$

(2) If $\lambda(I, A) \cap \lambda(F) = \emptyset$, then (3.1) has a unique solution

$$X = Q_c(A, C_1) S_r(I, A) Q_o(F, C_2, n) f_{(I,A)}^{-1}(F).$$

Proof. In view of relation (3.6), it is obvious that

$$\begin{bmatrix} R_0 C_1 & R_1 C_1 & \cdots & R_{n-1} C_1 \end{bmatrix} = Q_c(A, C_1) S_r(I, A).$$

Then we have

$$\begin{aligned} \sum_{i=0}^{n-1} R_i C F^i &= \begin{bmatrix} R_0 C_1 & R_1 C_1 & \cdots & R_{n-1} C_1 \end{bmatrix} Q_o(F, C_2, n) \\ &= Q_c(A, C_1) S_r(I, A) Q_o(F, C_2, n). \end{aligned}$$

By combining this relation with Theorem 2, we can obtain the conclusions. ■

4. Matrix equation $X - A\bar{X}F = C$

In this section, we consider the following matrix equation

$$X - A\bar{X}F = C, \quad (4.1)$$

by means of a real representation, where $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. We first define the real representation matrix equation of (4.1) by

$$Y - A_\sigma Y F_\sigma = C_\sigma. \quad (4.2)$$

According to the result in [14], we have the following results.

Proposition 1. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. Then matrix equation (4.1) has a solution X if and only if its real representation matrix equation (4.2) has a real matrix solution $Y = X_\sigma$.

Lemma 5. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. Then matrix equation (4.1) has a solution X if and only if matrix equation (4.2) has a solution $Y \in \mathbb{R}^{2n \times 2p}$, in which case, if Y is a solution to matrix equation (4.2), then the following matrix

$$X = \frac{1}{4} \begin{bmatrix} I_n & iI_n \end{bmatrix} (Y + Q_n Y Q_p) \begin{bmatrix} I_p \\ iI_p \end{bmatrix}$$

is a solution to matrix equation (4.1).

Based on the above results, we can obtain the solution of matrix equation (4.1) by solving its corresponding real representation matrix equation (4.2) according to the method proposed in the previous section. However, the dimensions of (4.2) are greater than the ones of (4.1). This may result in poor numerical stability. In the following theorem, a solution is established by the original matrices of the matrix equation, instead of the real representation matrices.

Theorem 5. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. Then,

(1) If X is a solution to matrix equation (4.1), then

$$X f_{(I, A\bar{A})}(\bar{F}F) = \sum_{k=0}^{n-1} \alpha_k \left(\sum_{j=k}^{n-1} (A\bar{A})^{j-k} C (\bar{F}F)^j + \sum_{j=k}^{n-1} (A\bar{A})^{j-k} (A\bar{C}F) (\bar{F}F)^j \right)$$

where $f_{(I, A\bar{A})}(s) = \alpha_n s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + \alpha_0$.

(2) If matrix equation (4.1) has a unique solution X , then

$$X = \sum_{k=0}^{n-1} \alpha_k \left(\sum_{j=k}^{n-1} (A\bar{A})^{j-k} C (\bar{F}F)^j + \sum_{j=k}^{n-1} (A\bar{A})^{j-k} (A\bar{C}F) (\bar{F}F)^j \right) f_{(I, A\bar{A})}^{-1}(\bar{F}F).$$

Proof. If matrix equation (4.1) has a solution X , then Eq. (4.2) has a solution $Y = X_\sigma$. By Theorem 3, we have

$$X_\sigma f_{(I, A_\sigma)}(F_\sigma) = \sum_{k=0}^{n-1} \sum_{s=2k}^{2n-1} \alpha_k A_\sigma^{s-2k} C_\sigma F_\sigma^s.$$

From Theorem 1 and Lemma 4, we have

$$\begin{aligned} (X f_{(I, A\bar{A})}(\bar{F}F))_\sigma &= X_\sigma \left(f_{(I, A\bar{A})}(\bar{F}F) \right)_\sigma P_p = X_\sigma (f_{(I, A\bar{A})}(\bar{F}F))_\sigma P_p = X_\sigma f_{(I, A_\sigma)}((F)_\sigma) \\ &= \sum_{k=0}^{n-1} \sum_{s=2k}^{2n-1} \alpha_k A_\sigma^{s-2k} C_\sigma F_\sigma^s \\ &= \sum_{k=0}^{n-1} \alpha_k \left(\sum_{j=k}^{n-1} A_\sigma^{2j-2k} C_\sigma (F)_\sigma^{2j} + \sum_{j=k}^{n-1} A_\sigma^{2j+1-2k} C_\sigma F_\sigma^{2j+1} \right) \\ &= \sum_{k=0}^{n-1} \alpha_k \left(\sum_{j=k}^{n-1} ((A\bar{A})^{j-k})_\sigma P_n C_\sigma ((F\bar{F})^j)_\sigma P_p + \sum_{j=k}^{n-1} ((A\bar{A})^{j-k})_\sigma P_n A_\sigma C_\sigma F_\sigma ((F\bar{F})^j)_\sigma P_p \right) \\ &= \sum_{k=0}^{n-1} \alpha_k \left(\sum_{j=k}^{n-1} ((A\bar{A})^{j-k} C (\bar{F}F)^j)_\sigma + \sum_{j=k}^{n-1} ((A\bar{A})^{j-k} (A\bar{C}F) (\bar{F}F)^j)_\sigma \right). \end{aligned}$$

So the first conclusion holds. With this the second conclusion is obviously true. ■

Remark 2. Different from the result in [14], the result given in the above theorem can be obtained from the original matrices, instead of the real representation of matrices.

Parallel to the results of matrix equation (3.1), we present some equivalent forms for the solution given in Theorem 5. On the basis of the above section, the proofs are very simple, and are omitted.

Theorem 6. Given $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$, let

$$f_{(I, A\bar{A})}(s) = \alpha_n s^n + \alpha_{n-1} s^{n-1} + \cdots + \alpha_1 s + \alpha_0, \quad \alpha_0 = 1$$

and

$$\text{adj}(I - sA\bar{A}) = R_{n-1} s^{n-1} + \cdots + R_1 s + R_0.$$

Then

(1) If X is a solution to matrix equation (4.1), then

$$X f_{(I, A\bar{A})}(\bar{F}F) = \sum_{i=0}^{n-1} R_i C (\bar{F}F)^i + \sum_{i=0}^{n-1} R_i (A\bar{C}F) (\bar{F}F)^i.$$

(2) If $\lambda(I, A\bar{A}) \cap \lambda(F\bar{F}) = \emptyset$, then matrix equation (4.1) has a unique solution

$$X = \left(\sum_{i=0}^{n-1} R_i C (\bar{F}F)^i + \sum_{i=0}^{n-1} R_i (A\bar{C}F) (\bar{F}F)^i \right) f_{(I, A\bar{A})}^{-1}(\bar{F}F).$$

Theorem 7. Let $A \in \mathbb{C}^{n \times n}$, $F \in \mathbb{C}^{p \times p}$, $C \in \mathbb{C}^{n \times p}$. Then

(1) If X is a solution to matrix equation (4.1), then

$$X f_{(I, A\bar{A})}(\bar{F}F) = Q_c(A\bar{A}, C) S_p(I, A\bar{A}) Q_o(\bar{F}F, I_p) + Q_c(A\bar{A}, A) S_n(I, A\bar{A}) Q_o(\bar{F}F, \bar{C}F).$$

(2) If $\lambda(I, A\bar{A}) \cap \lambda(F\bar{F}) = \emptyset$, then matrix equation (4.1) has a unique solution

$$X = (Q_c(A\bar{A}, C) S_p(I, A\bar{A}) Q_o(\bar{F}F, I_p) + Q_c(A\bar{A}, A) S_n(I, A\bar{A}) Q_o(\bar{F}F, \bar{C}F)) f_{(I, A\bar{A})}^{-1}(\bar{F}F).$$

5. Illustrative example

Example 1. Consider a matrix equation in the form of (4.1) with the following parameters:

$$A = \begin{bmatrix} 1 & -2-i & -1+i \\ 0 & i & 0 \\ 0 & -1 & 1-i \end{bmatrix}, \quad F = \begin{bmatrix} 2i & i \\ 1 & -1+i \end{bmatrix}, \quad C = \begin{bmatrix} -1+i & 1 \\ 0 & i \\ -i & 1-2i \end{bmatrix},$$

where $n = 3$, $p = 2$. It is easily checked that $\lambda(I, A\bar{A}) \cap \lambda(F\bar{F}) = \emptyset$, so this matrix equation has a unique solution. By some computations, we have

$$f_{(I, A\bar{A})}(s) = -2s^3 + 5s^2 - 4s + 1,$$

$$S_1(I, A\bar{A}) = \begin{bmatrix} 1 & -4 & 5 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix},$$

$$Q_c(A\bar{A}, C) S_p(I, A\bar{A}) Q_o(\bar{F}F, I_p) = \begin{bmatrix} -123 + 53i & -68 + 66i \\ -9 - 9i & -9 - 7i \\ -15 + 13i & -3 + 7i \end{bmatrix},$$

$$Q_c(A\bar{A}, A) S_n(I, A\bar{A}) Q_o(\bar{F}F, \bar{C}F) = \begin{bmatrix} 95 - 203i & 164 - 34i \\ 11 - 27i & 25 - 7i \\ 13 - 49i & 19 - 9i \end{bmatrix}.$$

Thus it follows from Theorem 7 that the unique solution of the matrix equation is

$$X = \begin{bmatrix} -\frac{877}{328} - \frac{745}{328}i & \frac{229}{328} - \frac{907}{328}i \\ -\frac{1}{4} - \frac{1}{2}i & \frac{1}{2} - \frac{3}{4}i \\ -\frac{69}{164} - \frac{23}{41}i & \frac{13}{41} - \frac{119}{164}i \end{bmatrix}.$$

Acknowledgements

This work was supported by the Program Changjing Scholars and Innovative Research Team in University, National Natural Science Foundation of China under Grant No. 60772046 and Natural Science Foundation of Guangdong Province under Grant No. 7301716.

Appendix. The proof of Theorem 1

Lemma 6. If $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times m}$, then $\det(I_m - AB) = \det(I_n - BA)$.

This lemma is standard. According to this lemma, the following result is readily obtained.

Lemma 7. For any $A \in \mathbb{C}^{n \times n}$, $f_{(I, A\bar{A})}(s) = f_{(I, \bar{A}A)}(s) \in \mathbb{R}[s]$.

Proof. It is obvious from the above lemma that $f_{(I, A\bar{A})}(s) = f_{(I, \bar{A}A)}(s)$. Thus we have $\lambda(I, A\bar{A}) = \lambda(I, \bar{A}A)$. On the other hand, if $\gamma \in \lambda(I, A\bar{A})$, then $\bar{\gamma} \in \lambda(I, \bar{A}A)$. By these two aspects, it follows that $f_{(I, A\bar{A})}(s) = f_{(I, \bar{A}A)}(s) \in \mathbb{R}[s]$. ■

Now we give another real representation of a complex matrix. It is a natural generalization of the following real representation of a complex number:

$$a + ib \rightsquigarrow \begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \quad a, b \in \mathbb{R}.$$

For distinction from the real representation given in Section 2.3, it is called the first real representation. The complete definition is given next.

Definition 2. For a complex matrix $A = A_1 + iA_2$, with $A_i \in \mathbb{R}^{m \times n}$, $i = 1, 2$, the first real representation $\bar{\sigma}$ is defined as

$$A_{\bar{\sigma}} = \begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix}.$$

Lemma 8. For any $A \in \mathbb{C}^{n \times n}$, if $\gamma \in \lambda(I, A)$, then $\{\gamma, \bar{\gamma}\} \subset \lambda(I, A_{\bar{\sigma}})$. Moreover, γ and $\bar{\gamma}$ have the same algebraic and geometric multiplicities.

Proof. Denote the algebraic and geometric multiplicities of γ by m and q , respectively. Further, denote the orders of q Jordan blocks associated with γ by p_i , $i = 1, 2, \dots, q$, and assume that the relative eigenvector chains associated with eigenvalue γ is $\{\eta_{ij}, i = 1, 2, \dots, q; j = 1, 2, \dots, p_i\}$. According to the definitions, we have

$$\eta_{ij} = (A_1 + iA_2)(\gamma \eta_{ij} + \eta_{i,j-1}), \quad \eta_{i,0} = 0. \quad (\text{A.1})$$

This relation is equivalent to

$$-i\eta_{ij} = (A_2 - iA_1)(\gamma \eta_{ij} + \eta_{i,j-1}). \quad (\text{A.2})$$

Combining (A.1) and (A.2), gives

$$\begin{bmatrix} \eta_{ij} \\ -i\eta_{ij} \end{bmatrix} = \begin{bmatrix} A_1 & -A_2 \\ A_2 & A_1 \end{bmatrix} \left(\gamma \begin{bmatrix} \eta_{ij} \\ -i\eta_{ij} \end{bmatrix} + \begin{bmatrix} \eta_{i,j-1} \\ -i\eta_{i,j-1} \end{bmatrix} \right).$$

This implies that the conclusion holds. ■

Lemma 9. Given $A \in \mathbb{C}^{n \times n}$, if $\gamma \in \lambda(I, (A_{\sigma})^2)$, then $\gamma \in \lambda(I, A\bar{A}) = \lambda(I, \bar{A}A)$.

Proof. Let $A = A_1 + iA_2$ as in Definition 5. then

$$(A\bar{A})_{\bar{\sigma}} = \begin{bmatrix} A_1^2 + A_2^2 & A_1A_2 - A_2A_1 \\ A_2A_1 - A_1A_2 & A_1^2 + A_2^2 \end{bmatrix} = (A_{\sigma})^2. \quad (\text{A.3})$$

In view of Lemmas 7 and 8, we have

$$\gamma \in \lambda(I, (A_{\sigma})^2) \implies \gamma \in \lambda(I, (A\bar{A})_{\bar{\sigma}}) \implies \lambda(I, A\bar{A}) = \lambda(I, \bar{A}A). \quad \blacksquare$$

The following lemma can be found in [14].

Lemma 10. Given $A \in \mathbb{C}^{n \times n}$, if $\gamma \in \lambda(I, A_{\sigma})$, then $\{\pm\gamma, \pm\bar{\gamma}\} \subset \lambda(I, A_{\sigma})$.

With the above preliminaries, we give the proof of Theorem 1.

Proof of Theorem 1. It follows from the above lemmas that

$$\gamma \in \lambda(I, A_{\sigma}) \implies \gamma^2 \in \lambda(I, (A_{\sigma})^2) \implies \gamma^2 \in \lambda(I, A\bar{A}) = \lambda(I, \bar{A}A).$$

On the other hand, by applying Lemma 8 it follows from (A.3) that the eigenvalue γ^2 of $(I, A\bar{A})$ (or $(I, \bar{A}A)$) and the eigenvalue γ of (I, A_{σ}) have common algebraic multiplicity. Combining this with the results of Lemmas 7 and 10, completes the proof. ■

References

- [1] R.H. Bartels, G.W. Stewart, A solution of the equation $AX + XB = C$, Commun. ACM 15 (1972) 820–826.
- [2] G. Kitagawa, An algorithm for solving the matrix equation $X = FXF^T + S$, Int. J. Control 25 (1977) 745–753.
- [3] D.L. Kleinman, P.K. Rao, Extensions to the Bartels–Stewart algorithm for linear matrix equation, IEEE Trans. Automat. Control 23 (1) (1978) 85–87.

- [4] G. Golub, S. Nash, C. Van Loan, A Hessenberg–Schur method for the problem $AX + XB = C$, IEEE Trans. Automat. Control 24 (1979) 909–913.
- [5] A. Betser, N. Cohen, E. Zeheb, On solving the Lyapunov and Stein equations for a companion matrix, System Control Lett. 25 (1995) 211–218.
- [6] E.C. Ma, A finite series solution of the matrix equation $AX - XB = C$, SIAM. J. Appl. Math. 14 (1966) 490–495.
- [7] R.W. Brockett, Introduction to Matrix Analysis, John Wiley, New York, 1970.
- [8] N.J. Young, Formulae for the solution of Lyapunov matrix equation, Internat. J. Control 31 (1980) 159–179.
- [9] E. DeSouza, S. Bhattacharyya, Controllability and observability and the solution of $AX - XB = C$, Linear Algebra Appl. 39 (1981) 167–188.
- [10] B. Zhou, G.R. Duan, An explicit solution to the matrix equation $AX - XF = BY$, Linear Algebra Appl. 402 (2005) 345–366.
- [11] J. Jones, J.R. Charleslew, Solutions of the Lyapunov matrix equation $BX - XA = C$, IEEE Trans. Automat. Control 27 (2) (1982) 464–466.
- [12] B. Hanzon, A faddeev sequence method for solving Lyapunov and Sylvester equations, Linear Algebra Appl. (1996) 401–430.
- [13] A. Jameson, Solution of the equation $AX - XB = C$ by inversion of an $M \times M$ or $N \times N$ matrix, SIAM J. Appl. Math. 16 (1968) 1020–1023.
- [14] T. Jiang, M. Wei, On solutions of the matrix equations $X - AXB = C$ and $X - A\bar{X}B = C$, Linear Algebra Appl. 367 (2003) 225–233.
- [15] A.G. Wu, G.R. Duan, Solution to the generalized Sylvester matrix equation $AV + BW = EVF$, IET Control Theory Appl. 1 (1) (2007) 402–408.
- [16] A.G. Wu, Y.M. Fu, G.R. Duan, On solutions of the matrix equations $V - AVF = BW$ and $V - A\bar{V}F = BW$, Math. Comput. Modelling 47 (2008) 1181–1197.
- [17] F.L. Lewis, Further remarks on the Cayley–Hamilton theorem and Leverrier’s method for matrix pencil $(sE - A)$, IEEE Trans. Automat. Control 31 (9) (1986) 869–870.